# Probability density estimation on the hyperbolic space applied to radar processing

Emmanuel Chevallier, Frédéric Barbaresco, Jesus Angulo

# Probability density estimation on the hyperbolic space applied to radar processing

Emmanuel Chevallier[1], Frédéric Barbaresco[2], and Jesús Angulo[1]

[1] MINES ParisTech, PSL-Research University,
CMM-Centre de Morphologie Mathématique, France
[2] Thales Air Systems, Surface Radar Domain, Technical Directorate,
Advanced Developments Department, 91470 Limours, France
emmanuel.chevallier@mines-paristech.fr

**Abstract.** Main techniques of probability density estimation on Riemannian manifolds are reviewed in the hyperbolic case. For computational reasons we chose to focus on the kernel density estimation and we provide the expression of Pelletier estimator on hyperbolic space. The method is applied to density estimation of reflection coefficients from radar observations.

## 1 Introduction

The problem of probability density estimation is a vast topic. Their exists several standard methods in the Euclidean context, such as histograms, kernel methods, or the characteristic function method. These methods can sometimes be transposed to the case of Riemannian manifolds. However, the transposition often introduces additional computational efforts. This additional effort depends on the method used and the nature of the manifold. The hyperbolic space is one of the most elementary non-Euclidean spaces. It is one of the three simply connected isotropic manifolds, the two others being the sphere and the Euclidean space. The specificity of the hyperbolic space enables to adapt the different density estimation methods at a reasonable cost. Convergence rates of the density estimation using kernels and orthogonal series were progressively generalized to Riemannian manifolds, see [13][14][16]. More recently convergence rates for the kernel density estimation without the compact assumption have been introduced [7], which enables the use of Gaussian-type kernels [?]. One already encounters the problem of density estimation in the hyperbolic space for electrical impedance [14] and networks [8]. We are interested here in the estimation of the density of the reflection coefficients extracted from a radar signal [3]. These coefficients have a intrinsic hyperbolic structure [4,2]. For computational reasons we chose to focus our applications on the kernel density estimation. The paper begins with an introduction to the hyperbolic geometry in Section 2. Section 3 reviews the two main density estimation techniques on the hyperbolic space. Section 4 presents an application to radar data estimation.

## 2 The hyperbolic space and the Poincaré disk model

The hyperbolic geometry results of a modification of the fifth Euclid's postulate on parallel lines. In two dimensions, given a line $D$ and a point $p \notin D$, the hyperbolic geometry is an example where there are at least two lines going through $p$, which do not intersect $D$. Let us consider the unit disk of the Euclidean plane endowed with the Riemannian metric:

$$ds_{\mathbb{D}}^2 = 4\frac{dx^2 + dy^2}{(1 - x^2 - y^2)^2} \tag{1}$$

where $x$ and $y$ are the Cartesian coordinates. The unit disk $\mathbb{D}$ endowed with $ds_{\mathbb{D}}$ is called the Poincaré disk and is a model of the two-dimensional hyperbolic geometry $\mathcal{H}_2$. The construction is generalized to higher dimensions. It can be shown that the obtained Riemannian manifold $\mathcal{H}_n$ is homogeneous. In other words,

$$\forall p, q \in \mathcal{H}_n, \exists \varphi \in ISO(\mathcal{H}_n), \varphi(p) = q$$

where $ISO(\mathcal{H}_n)$ is the set of isometric transformations of $\mathcal{H}_n$.

In $\mathbb{R}^n$ the convolution of a function $f$ by a kernel $g$ consists in the integral of translated kernel weighted by $f$ in each point $p$ of the support space. The group law $+$ of $\mathbb{R}^n$ is an isometry that enables to transport the kernel in the whole space. In the Riemannian setting the definition of convolution $f * g$ needs some homogeneity assumption: in an homogeneous space it is possible to transport a kernel from a reference point to any other point by an isometry. Let the isotropy group of $p$ be the set of isometries that fix $p$. The convolution is properly defined for kernels that are invariant under elements of the isotropy group of $p$, for a $p$ in $\mathcal{H}_n$. Formally, let $K_{p_{ref}}$ be a function invariant to the isotropy group of $p_{ref}$. Let $K_p = K_{p_{ref}} \circ \varphi_{p,p_{ref}}$ with $\varphi_{p,p_{ref}}$ an isometry such that $\varphi_{p,p_{ref}}(p) = p_{ref}$. When it exists, since $\mathcal{H}_n$ is homogeneous, we can define the convolution of a function $f$ by the kernel $K_{p_{ref}}$ by:

$$(f * K_{p_{ref}})(q) = \int f(p) K_p(q) dvol$$

where $vol$ is the hyperbolic measure.

Furthermore it can be shown that for any couple of geodesic $\gamma_1$ and $\gamma_2$ starting from $p \in \mathcal{H}_n$, there exists $\varphi$ in the isotropy group of $p$ such that $\varphi(\gamma_1) = \gamma_2$. In other words, a kernel $K_p$ invariant under the isotropy group of $p$ looks the same in every directions at point $p$. For more details on the hyperbolic space, see [6].

## 3 Non-parametric probability density estimation

Let $\Omega$ be a space endowed with a probability measure. Let $X$ be a random variable $\Omega \rightarrow \mathcal{H}_n$. The measure on $\mathcal{H}_n$ induced by $X$ is noted $\mu_X$. We assume that $\mu_X$ has a density, noted $f$, with respect to $vol$, and that the support of $X$ is a compact set noted $Supp(X)$. Let $(x_1, .., x_k) \in (\mathcal{H}_n)^k$ be a set of draws of $X$.

Let $\mu_k = \frac{1}{k}\sum_i \delta_{x_i}$ denote the empirical measure of the set of draws. This section presents the main techniques of estimation of $f$ from the set of draws $(x_1, .., x_k)$. The estimated density at $x \in \mathcal{H}_n$ is noted $\hat{f}_k(x) = \hat{f}_k(x, x_1, ..., x_k)$. Observe that $\hat{f}_k(x)$ can be seen as a random variable. The relevance of density estimation technique depends on several aspects. Recall that $\mathcal{H}_n$ is isotropic, every points and directions are indiscernible. In absence of prior information on the density, the estimation technique should not privilege specific directions or locations. This results in a homogeneity and an isotropy condition. The convergence of the different estimation techniques is widely studied. Results were first obtained in the Euclidean case, and are progressively extended to the probability densities on manifold [16][13][14][7]. The last aspect, is computational. Each estimation technique has its own computational framework, which presents pro and cons given the different applications. For instance, the estimation by orthogonal series presents an initial pre-processing, but provides a fast evaluation of the estimated density in compact manifolds. These aspects are studied for the main techniques of density estimation in the remaining of the paper.

Every standard density estimation technique involves a scaling parameter. This scaling factor controls the influence of the observation $x_i$ on the estimated density at $x$, depending on the distance between $x$ and $x_i$. In the experiments, the scaling factor is chosen following the framework proposed in [9]: a cross validation of the likelihood of the estimator.

### 3.1 Characteristic function, or orthogonal series

Let $U \in \mathcal{H}_n$ be such that $Supp \subset U$. Let the functions $(e_j)$ be the eigenfunctions $(e_j)$ of the Laplace operator on $L^2(U)$. Recall that

$$\langle e_i, e_j \rangle = \int_U e_i \overline{e_j} dvol,$$

where $\overline{x}$ denotes the complex conjugate of $x$. Eigenfunctions of the Laplace operator in $\mathcal{H}_n$ behave similarly to the eigenfunctions in the real case. Thus, if the density $f$ lays in $L^2(U)$ the Fourier-Helgason transform gives:

$$f = \sum_j \langle f, e_j \rangle e_j, \text{ or } f = \int_{\mathcal{B}} \langle f, e_j \rangle e_j de_j,$$

respectively when $U$ is compact and non compact. See [11] for an expression of $de_j$ and $\mathcal{B}$ when $U = \mathcal{H}_n$. By the law of large number, one has

$$\langle f, e_i \rangle = \int f \overline{e_j} dvol = \mathbb{E}\left(\overline{e_j}(X)\right) \approx \frac{1}{k}\sum_{j=1}^{k} \overline{e_j}\left(x_i\right).$$

Given $T > 0$, let $\mathcal{B}_T \subset \mathcal{B}$. The density estimator becomes:

$$\hat{f}_k = \frac{1}{k}\sum_{e_j \in \mathcal{B}_T}\left[\sum_{i=1}^{k} \overline{e_j}\left(x_i\right)\right] e_j, \text{ or } \hat{f}_k = \frac{1}{k}\int_{\mathcal{B}_T}\left[\sum_{i=1}^{k} \overline{e_j}\left(x_i\right)\right] e_j de_j.$$

For a suitable choice of $\mathcal{B}_T$, see [14][13], the parameter $T$ plays the role of the inverse of the scaling parameter. In the Euclidean context, this is equivalent to the characteristic function density estimator. The choice of the basis is motivated by the regularity of the eigenfunctions of the Laplace operator. Let $I_T$ be the indicator function of $\mathcal{B}_T$ in $\mathcal{B}$. For a right $\mathcal{B}_T$, $\mathcal{FH}^{-1}(I_T)$ is invariant under the isotropy group of a $p \in \mathcal{H}_n$. Then, the estimation is a convolution [12] written as

$$\hat{f}_k = \mu_k * \mathcal{FH}^{-1}(I_T),$$

where $\mathcal{FH}$ is the Fourier Helgason transform. In other words, the estimation does not privileges specific locations or directions. Convergence rates are provided in [14]. When $U$ is compact, the estimation $\hat{f}$ is made through the estimation of $N$ scalar product, that is to say $kN$ summation operation. However, the evaluation at $x \in \mathcal{H}_n$ involves only a sum of $N$ terms. On the other hand, when $U$ is not compact, the evaluation of the integral requires significantly higher computational cost. Unfortunately, the eigenfunction of the Laplacian for $U \subset \mathcal{H}_n$ are only known for $U = \mathcal{H}_n$. See [15] and [13] for more details on orthogonal series density estimation on Riemannian manifolds.

### 3.2 Kernel density estimation

Let $K : \mathbb{R}_+ \to \mathbb{R}_+$ be a map which verifies the following properties: $\int_{\mathbb{R}^n} K(||x||)dx = 1$, $\int_{\mathbb{R}^n} xK(||x||)dx = 0$, $K(x > 1) = 0$, $sup(K(x)) = K(0)$. Given a point $p \in \mathcal{H}_n$, $exp_p$ defines a new injective parametrization of $\mathcal{H}_n$. The Lebesgue measure of the tangent space is noted $Leb_p$. The function $\theta_p : \mathcal{H}_n \to \mathbb{R}_+$ defined by:

$$\theta_p : q \mapsto \theta_p(q) = \frac{dvol}{dexp^*(Leb_p)}(q), \tag{2}$$

is the density of the Riemannian measure with respect to the image of the Lebesgue measure of $T_p\mathcal{H}_n$ by $exp_p$. Given $K$ and a scaling parameter $\lambda$, the estimator of $f$ proposed by Pelletier in [16] is defined by:

$$\hat{f}_k = \frac{1}{k} \sum_i \frac{1}{\lambda^n} \frac{1}{\theta_{x_i}(x)} K\left(\frac{d(x, x_i)}{\lambda}\right). \tag{3}$$

It can be noted that this estimator is the usual kernel estimator is the case of Euclidean space. Convergence rates are provided in [16]. These rates are similar to those of the orthogonal series. In [16] is furthermore shown that $x_i$ is the intrinsic mean of $\frac{1}{\theta_{x_i}(.)} K(d(., x_i)/\lambda)$. Under reasonable assumptions on the true density $f$, the shape of the kernel does not have a significant impact on the quality of the estimation in the Euclidean context [17]. Fig.3.2 experimentally confirms the result in $\mathcal{H}_2$.

Given a reference point $p_{ref} \in \mathcal{H}_n$, let

$$\tilde{K}(q) = \frac{1}{k\lambda^n} \frac{1}{\theta_{p_{ref}}(q)} K\left(\frac{d(p_{ref}, q)}{\lambda}\right). \tag{4}$$

Note first that if $\phi$ is an isometry of $\mathcal{H}_n$, $\theta_p(q) = \theta_{\phi(p)}(\phi(q))$. After noticing that $\tilde{K}$ is invariant under is isotropy group of $p_{ref}$, a few calculations lead to:

$$\hat{f}_k = \mu_k * \tilde{K}. \tag{5}$$

As the density estimator based on the eigenfunction of the Laplacian operator, the kernel density estimator is a convolution and does not privileges specific locations or directions. In order to evaluate the estimated density at $x \in \mathcal{H}_n$, one first need to determine the observations $x_i$ such that $d(x, x_i) < \lambda$, and perform a sum over the selected observations.

One still needs to obtain an explicit expression of $\theta_p$. Given a reference point $p$, the point of polar coordinates $(r, \alpha)$ of the hyperbolic space is defined as the point at distance $r$ of $p$ on the geodesic with initial direction $\alpha \in \mathbb{S}^{n-1}$. Since the $\mathcal{H}_n$ is isotropic the expression the length element in polar coordinates depends only on $r$. Expressed in polar coordinates the hyperbolic metric expression is [1,10]:

$$\mathbf{g}_{\mathcal{H}_n} = dr^2 + sinh(r)^2 \mathbf{g}_{\mathbb{S}^{n-1}}.$$

The polar coordinates are a polar expression of the exponential map at $p$. In an adapted orthonormal basis of the tangent plane the metric takes then the following form:

$$G = \begin{pmatrix} 1 & 0 \\ 0 & sinh(r)^2 \frac{1}{r^2} I_{n-1} \end{pmatrix} \tag{6}$$

where $I_{n-1}$ is the identity matrix of size $n-1$. Thus, using (6), the volume element $d\mu_{exp_p^*}$ is given by

$$dvol = \sqrt{G}.dexp_p^*(Leb_p) = (\tfrac{1}{r} sinh(r))^{n-1} dexp_p^*(Leb_p). \tag{7}$$

where $r = d(p, q)$. From (2) and (7), one obtains

$$\theta_p(q) = (\tfrac{1}{r} sinh(r))^{n-1}. \tag{8}$$

Finally, plugging (8) into (3) gives

$$\hat{f}_k = \frac{1}{k} \sum_i \frac{1}{\lambda^n} \frac{d(x, x_i)^{n-1}}{sinh(d(x, x_i))^{n-1}} K\left(\frac{d(x, x_i)}{\lambda}\right). \tag{9}$$

## 4 Application to radar estimation

### 4.1 From radar observations to reflection coefficients $\boldsymbol{\mu_k} \in \mathbb{D}$

Let us discuss briefly how radar data is related to hyperbolic space via reflection coefficients, for more details see [2]. Each radar cell is a complex vector $z = (z_0, \cdots, z_{n-1})$ considered as a realization of a centered stationary Gaussian process $Z = (Z_0, \cdots, Z_{n-1})$ of covariance matrix $R_n = \mathbb{E}[ZZ^*]$. The matrix $R_n$ has a Toeplitz structure. For $1 \leq k \leq l \leq n-1$, the $k$-th order autoregressive

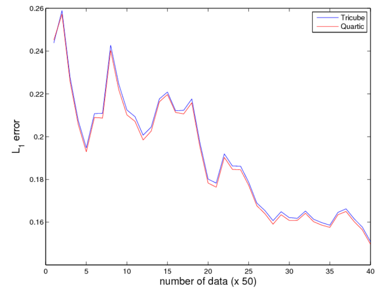


**Fig. 1.** Consider a law $X$ whose density in the tangent plane at $(0,0)$ is a centered hemisphere. From a set of draws, the density is estimated using two standard kernels, $K(x) = (1-x^2)^2\mathbf{1}_{x<1}$ and $K(x) = (1-x^3)^3\mathbf{1}_{x<1}$. The $L_1$ distance between the estimated and the true density is plotted depending on the number of draws.

estimate of $Z_l$ is given by $\hat{Z}_l = -\sum_{j=1}^{k} a_j^{(k)} Z_{l-j}$, where the autoregressive coefficients $a_1^{(k)} \cdots a_k^{(k)}$ are chosen such that the mean squared error $\mathbb{E}(|Z_l - \hat{Z}_l|^2)$ is minimized. In practice, reflection coefficients are estimated by regularized Burg algorithm [3]. The last autoregressive coefficient $a_k^{(k)}$ is called the $k$-th reflection coefficient, denoted by $\mu_k$ and which has the property $|\mu_k| < 1$. The coefficient for $k=0$ corresponds to the power, denoted $P_0 \in \mathbb{R}_+^*$. The reflection coefficients induce a (diffeomorphic) map $\varphi$ between the Toeplitz Hermitian positive definite (HPD) matrices of order $n$, $\mathcal{T}^n$, and reflection coefficients:

$$\varphi : \ \mathcal{T}^n \to \mathbb{R}_+^* \times \mathbb{D}^{n-1}, \ \ R_n \mapsto (P_0, \mu_1, \cdots, \mu_{n-1})$$

where $\mathbb{D} = \{\zeta \in \mathbb{C} \ : \ |\zeta| < 1\}$ is the open unit disk of the complex plane. Diffeomorphism $\varphi$ is very closely related to theorems of Trench [18].

The Riemannian geometry of the space of reflection coefficients has been explored in [4] through the Hessian of Kähler potential, whose metric is

$$ds^2 = n\frac{dP_0^2}{P_0^2} + \sum_{k=1}^{n-1}(n-k)\frac{|d\mu_k|^2}{(1-|\mu_k|^2)^2}. \tag{10}$$

According to the metric (10) the space $\mathcal{T}^n$ can be seen as a product of the Riemannian manifold $(\mathbb{R}_+^*, ds_0^2)$, with $ds_0^2 = n(dP_0^2/P_0^2)$ (logarithmic metric multiplied by $n$), and $(n-1)$ copies of $(\mathbb{D}, ds_k^2)_{1\le k\le n-1}$, with $ds_k^2 = (n-k)ds_{\mathbb{D}}^2$. $(\mathbb{R}_+^* \times \mathbb{D}^{n-1}, ds^2)$ is a Cartan–Hadamard manifold whose sectional curvatures are bounded, i.e., $-4 \le K \le 0$. This metric is related information geometry and divergence functions [5]. From the product metric, closed forms of the Riemannian distance, arc-length parameterized geodesic, etc. can be obtained [4,2].

The next paragraph presents the estimations of the marginal densities coefficients $\mu_k$.

### 4.2 Experimental results

Data used in the experimental tests are radar observations from THALES X-band Radar, recorded during 2014 field trials campaign at Toulouse Blagnac Airport for European FP7 UFO study (Ultra-Fast wind sensOrs for wake-vortex hazards mitigation). Data are representative of Turbulent atmosphere monitored by radar in rainy conditions. Fig. 2 illustrates the density estimation of the six reflection coefficients on the Poincaré unit disk. For each coefficient the dataset is composed of 120 draws.

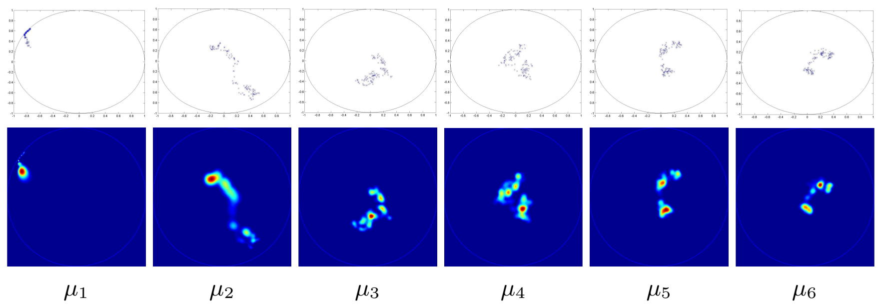


**Fig. 2.** Estimation of the density of the three first coefficients $\mu_k$ under rainy conditions. The expression of the used kernel is $K(x) = \frac{3}{\pi}(1-x^2)^2\mathbf{1}_{x<1}$

## 5 Conclusion and perspectives

We have discussed the problem of density estimation on the hyperbolic space. After having computed the volume change factor, we have adopted the approach based on kernel density estimation by Pelletier [16]. The method has been used to estimate the density of reflection coefficients from radar signals. According to the diffeomorphism between Toeplitz HPD matrices and reflection coefficients [18], densities estimated on a product of Poincaré disk can be interpreted as a probability density on the space of Toeplitz HPD matrces. Other symmetric homogenous spaces such as the Siegel disk can be addressed using similar methods. The link between the Siegel disk and the space of Toeplitz-Block-Toeplitz HPD matrices makes of it an other interesting study case.

Alternative approaches of density estimation can be considered in future research. For data lying in a known compact symmetric subspace of the hyperbolic space, it is possible to use the orthogonal series technique, where the eigenfunctions of Laplace operator are numerically estimated. From an application viewpoint, densities from radar reflection coefficients can be used as basic ingredient in radar detection algorithms (finding modes of density and segmenting the density).

## References


1. J.-Ph. Anker, P. Ostellari. The heat kernel on noncompact symmetric spaces. In *Lie groups and symmetric spaces,* AMS Transl. Ser. 2, 210, 27Ű-46, 2003.
2. M. Arnaudon, F. Barbaresco, L. Yang. Riemannian Medians and Means With Applications to Radar Signal Processing. IEEE Journal of Selected Topics in Signal Processing, 7(4): 595–604, 2013.
3. F. Barbaresco. Super resolution spectrum analysis regularization: Burg, capon and ago antagonistic algorithms. In *Proc. of EUSIPCO'96,* pp. 2005-2008, Trieste, 1996.
4. F. Barbaresco. Information Geometry of Covariance Matrix: Cartan-Siegel Homogeneous Bounded Domains, Mostow/Berger Fibration and Fréchet Median. In *(F. Nielsen, R. Bhatia, Eds.) Matrix Information Geometry,* Springer, pp. 199–255, 2013.
5. F. Barbaresco. Koszul Information Geometry and Souriau Geometric Temperature/Capacity of Lie Group Thermodynamics. *Entropy,* 16:4521–4565, 2014
6. J.W. Cannon, W.J. Floyd, R. Kenyon, W.R. Parry. *Hyperbolic Geometry.* Flavors of Geometry, MSRI Publications, Vol. 31, 1997.
7. D. Asta. Kernel Density Estimation on Symmetric Spaces. *arXiv preprint arXiv:1411.4040,* 2014.
8. D. Asta, C. Shalizi. Geometric network comparison. *arXiv preprint arXiv:1411.1350,* 2014.
9. R.P.W. Duin. On the choice of smoothing parameters for Parzen estimators of probability density functions. *IEEE Transactions on Computers,* 25(11): 1175–1179, 1976.
10. A. Grigor'yan. *Heat kernel and analysis on manifolds*, American mathematical soc., 47, 2012.
11. S. Helgason. *Non-Euclidean Analysis*, arXiv:math/0411411 [math.DG], 2004.
12. S. Helgason. The Abel, Fourier and Radon transform on symmetric spaces. *Indagationes Mathematicae,* 16(3):531–551, 2005.
13. H. Hendriks. *Nonparametric estimation of a probability density on a Riemannian manifold using Fourier expansions.* The Annals of Statistics, 18(2): 832–849, 1990.
14. S. Huckeman, P. Kim, J. Koo, A. Munk. Mobius deconvolution on the hyperbolic plan with application to impedance density estimation. *The Annals of Statistics,* 38(4):2465–2498, 2010.
15. P. Kim, D. Richards. Deconvolution density estimation on the space of positive definite symmetric matrices. In *Nonparametric Statistics and Mixture Models: A Festschrift in Honor of Thomas P. Hettmansperger,* pp. 147–168, World Scientific Publishing, 2008.
16. B. Pelletier. Kernel density estimation on Riemannian manifolds. *Statistics and Probability letters,* 73(3):297–304, 2005.
17. B.W. Silverman. *Density estimation for statistics and data analysis*, Chapman and Hall, 1986.
18. W.F. Trench. An algorithm for the inversion of finite Toeplitz matrices. *J. Soc. Indust. Appl. Math.,* 12: 515–522, 1964.